An Improved Procedure for VaR/CVaR Estimation under Stochastic Volatility Models

Chuan-Hsiang Han\textsuperscript{1}  
Wei-Han Liu\textsuperscript{2}  
Tzu-Ying Chen\textsuperscript{3}

This version: June 29, 2010

Abstract: This paper proposes an improved procedure for stochastic volatility model estimation with an application in risk management. This procedure is composed of the following instrumental components: Fourier transform method for volatility estimation with a price correction scheme, and importance sampling for extremal event probability estimation with applications to estimate Value-at-Risk and conditional Value-at-Risk. Then we conduct a Value-at-Risk backtesting for some foreign exchange data and the S&P 500 index data. In comparison with empirical results obtained from RiskMetrics, historical simulation, and the GARCH(1,1) model, we find that our improved procedure outperforms on average.

JEL classification: C13; C14; C63.

Keywords: stochastic volatility, Fourier transform method, importance sampling, (conditional) Value-at-Risk, backtesting.

\textsuperscript{1} Department of Quantitative Finance, National Tsing-Hua University, Hsinchu,Taiwan 30013, ROC. E-mail: chhan@mx.nthu.edu.tw. Work supported by NSC 97-2115-M-007-002-MY2, Taiwan.
\textsuperscript{2} School of Economics and Finance, La Trobe University, Melbourne, Australia. E-mail: weihanliu2002@yahoo.com.
\textsuperscript{3} Department of Finance, National Taiwan University.
1 Introduction

There are two major approaches for Value-at-Risk (VaR) and conditional Value-at-Risk (CVaR) estimation, two of the most popular risk measures: modeling the return distribution and capturing the volatility process (Jorion, 2007). For the former, various techniques are employed for modeling the whole return distribution or just the tail areas, including known parametric distribution, kernel density approximation, and extreme value theory, etc. The latter heavily relies on discrete-time volatility models such as EWMA and GARCH to capture the volatility process. See Jondeau et al. (2007) for further details.

Continuous-time stochastic volatility models are known for demonstrating some stylized features of financial data and they have been intensively applied in option pricing and hedging. Fouque et al. (2000) derive option pricing and hedging approximation formula under these models by means of a singular perturbation technique. Lehar et al. (2002) take option data to calibrate a stochastic volatility model under an equivalent martingale (risk-neutral) probability measure, and use VaR as a performance criterion to compare with other pricing models. They incorporate an option pricing approximation for model calibration and use empirical percentile estimated from simulated option price change to approximate VaR. These approximations are crucial to reduce computational cost.

Our main interest in this paper is to investigate estimation and performance of stochastic volatility models under the historical (or physical) probability measure with applications in risk management. We primarily work on spot price data such as foreign exchange rates and S&P 500 index, instead of option data and its pricing mechanism, for risk measure estimation. In addition, we value exact computation by means of efficient simulation methods. There is no need to incorporate approximation results as discussed in previous studies so that approximation errors can be avoided. In the direction of our investigation subject, major hurdles include (i) unstable parameter estimation subject to practical data constrain, and (ii) lack of efficient computational technique for accurately estimating risk measures. This paper provides a procedure for VaR and CVaR estimation under stochastic volatility models by proposing improvements in the following two aspects: (i) improved stochastic volatility model estimation scheme by refining Fourier transform method (Malliavin and Mancino (2002, 2009)) with a price correction, and (ii) enhanced importance sampling for estimating extremal event probability.

For the first aspect, we proceed to propose improved methods for stochastic volatility estimation. Stochastic volatility models have wide applications and their estimation issues present interesting challenges. Various estimation procedures have been proposed for stochastic volatility model parameter estimation. Broto and Ruiz (2004) provide an overview of stochastic volatility estimation methods including method of moments, generalized method of moments, maximum likelihood estimators, quasi maximum likelihood, etc. Yu (2010) emphasizes simulation-based estimation methods. Simulated maximum likelihood, simulated generalized method of moments, efficient method of moments, indirect inference and Markov chain Monte Carlo, etc are reviewed. All these aforementioned estimation methods are parametric.
Among non-parametric approaches to estimate the integrated volatility, many recent literatures exploit quadratic variation formula. See Zhang et al. (2005) and references therein. Using integrated volatility to approximate instantaneous or spot volatility is possibly infeasible because its differentiation procedures may be numerically unstable and its modeling performance varies with data frequency, as cautioned by Malliavin and Mancino (2009). The authors propose a non-parametric Fourier transform method to estimate spot volatility under continuous semi-martingale models. This method relies on the integration of the time series rather than on its differentiation because it is based on the computation of Fourier coefficients of the variance process rather than on quadratic variation. The authors also claim that this approach is particularly suitable for the analysis of high frequency time series and for the computation of cross volatilities. We adopt this Fourier transform method as the framework for stochastic volatility model estimation.

However, Reno (2008) alerts that the Fourier algorithm performs badly near boundaries of estimated volatility time series data, i.e. estimated volatility of the first and last 1% time series are not accurate enough. The author recommends discarding those volatility estimates near the boundary. Yet, this compromise may constitute a drawback in estimation, more than a cost of information. An example of this drawback is that, when exclusively following Reno (2008), dropping the most recent 1% volatility estimates will distort the prediction of a short-time volatility, say one-day volatility. To avoid this “boundary effect” pitfall, we provide a price correction scheme by matching the estimated volatility with observed price returns. This scheme only requires solving a regression equation derived from the maximum likelihood method so it is easy to implement. Additional advantages include (i) no loss of data observations and (ii) reduction of the volatility bias generated from the Fourier transform method. In our simulation study, both mean squared error and maximum absolute error are reduced at least by half. Then we apply this corrected Fourier transform method to stochastic volatility model parameter estimation. To be specific, we first use the Fourier transform method with our price correction scheme to produce a volatility time series. Second, we apply the maximum likelihood method to estimate stochastic volatility model parameters. When model parameters are fully estimated, we proceed to compute quantities of interest: VaR and CVaR.

For the second aspect for VaR and CVaR estimation under stochastic volatility models, we propose an enhanced version of importance sampling for estimating extremal event probability. We aim to resolve a computational issue under stochastic volatility models for VaR/CVaR estimation. In general, there is no closed-form solution for these quantities under stochastic volatility models. We develop an efficient simulation method, namely importance sampling, as our computational tool. Notably, importance sampling is one of the key techniques for variance reduction in Monte Carlo simulation (Glasserman 2003, Lemieux 2009). It can effectively improve convergence of sample means particularly in rare event simulation. The theoretical background of our proposed importance sampling relies on a combination of the large deviation theory (Bucklew 2004) with the averaging effect of realized variance (Fouque et al. 2000). This methodology is useful to handle heavy (or fatter) tail distributions induced by stochastic volatility models.

Empirical analysis confirms the outperformance of VaR estimation by our proposed improved procedure including corrected Fourier transform method and importance sampling. Two datasets are used for empirical
examination: the first one contains three foreign exchange series (January 5, 1998 to July 24, 2009) and the second one contains S&P500 index and its VIX (January 3, 2005 to July 24, 2009). Data period covers both tranquil and turbulent times. Three types of backtesting are conducted for model evaluation and performance comparison in VaR estimation. Our proposed procedure significantly outperforms especially at 99% VaR estimates, as compared with RiskMetrics, historical simulation, and GARCH(1,1) model. This outperformance matches some demands from the Basel II Accord implementation (Jorion 2007), e.g. risk measurement at 99% significance level.

The organization of this paper is as follows. Section 2 introduces a general one-factor stochastic volatility model, the extremal event probability estimation, and their relationship with VaR and CVaR estimation. Section 3 reviews the Fourier transform method, a nonparametric approach to estimate volatility time series, and explores a price correction scheme to alleviate the “boundary effect.” Section 4 discusses in detail on construction of the efficient importance sampling estimators for extremal event probabilities, then solve for VaR/CVaR. Section 5 investigates backtesting results of VaR estimation over three foreign exchange rate series and S&P500 index with its VIX, and compare these results with some well known methods such as RiskMetrics, historical simulation, and GARCH(1,1). Section 6 concludes.

2 VaR/CVaR Estimation under Stochastic Volatility Models

The Black-Scholes model is fundamental in option pricing theory under no-arbitrage condition (Hull 2008). It simply assumes that log returns of risky asset prices are normally distributed. A stochastic volatility model is an extension of the Black-Scholes model by relaxing the assumption of constant volatility to allow volatility being driven by other processes. Under a probability space \( (\Omega, F, P, (F_t)_{0 \leq t \leq T}) \), a general form of one-factor stochastic volatility model is defined by

\[
\begin{align*}
    dS_t &= \mu S_t dt + \sigma_t S_t W_t^{(0)} , \\
    \sigma_t &= f(Y_t) , \\
    dY_t &= c(Y_t) dt + g(Y_t) \left( \rho dW_t^{(0)} + \sqrt{1 - \rho^2} dW_t^{(1)} \right) ,
\end{align*}
\]

where \( S_t \) denotes the underlying asset price process with a constant growth rate \( \mu \) and a random volatility \( \sigma_t \). The vector \( (W_t^{(0)}, W_t^{(1)}) \) consists of two independent standard Brownian motions and \( \rho \) denotes the instantaneous correlation coefficient satisfying \( |\rho| \leq 1 \). We further assume that the volatility function \( f \) and coefficient functions of \( Y_t \), namely \( (c(y), g(y)) \), satisfy classical assumptions such that the whole dynamic system (1) fulfills the existence and uniqueness conditions for the strong solution of stochastic differential equations (Oksendal 1998). The stochastic volatility model (1) is one-factor because its random volatility \( \sigma_t \) is driven by a single factor process \( Y_t \), also known as the driving volatility process. This process is often assumed mean-reverting. Typical examples include Ornstein-Uhlenbeck (OU) process and square-root process (Heston 1993). For these two processes, their coefficient functions \( (c(y), g(y)) \) are \( (\alpha(m-y), \beta) \) and \( (\alpha(m-y), \beta \sqrt{y}) \), respectively, where \( y \) denotes the variable of driving volatility. Other model parameters \( (\alpha, m, \beta) \) denote the rate of mean reversion, long-run mean, and the volatility of volatility, respectively. The volatility function \( f(y) \) are customarily chosen as \( \exp(y/2) \) and \( \sqrt{y} \), respectively.

In financial applications such as option pricing, hedging, and risk management, one often needs to compute...
the following conditional expectation under model (1) given a Markovian assumption:
\[ P(t, x, y) = E[e^{-r(t-t)} H(S_t) \mid S_0 = x, Y_t = y], \]  
(2)
where the value function \( P \) may represent an option price or a hedging ratio given state variables of asset price \( x \) and its driving volatility \( y \) at time \( t \geq 0 \). Other parameters or variables in (2) include the discounting rate \( r \), the payoff function \( H \), and the exercise time \( \tau(\geq t) \), that can be either a fixed maturity, say \( \tau = T \), or a stopping time. As an example of option pricing, Equation (2) has to be evaluated under an equivalent martingale measure so that the no-arbitrage condition holds. As an example in risk management, Equation (2) can be used to depict the tail areas, i.e. specific probability levels for extremal events under the historical or physical probability measure.

For our purpose to estimate VaR and CVaR, we shall first consider the estimation of an extremal event probability under the general formulation in (2). We do this by choosing the discounting rate \( r = 0 \), the payoff function as a rare-event indicator \( H(x) = \mathbf{1}(x < D) \), where \( D \) denotes a threshold or cutoff point, and the exercise time \( \tau = T > 0 \) a fixed date. Hence, the time-\( T \) probability of an extremal event for logarithmic returns conditional at time 0 is defined by
\[ P(0, x, y; D) = E[\mathbf{1}(\ln(S_T / S_0) < D) \mid S_0 = x, Y_0 = y]. \]
(3)
There are three major cases where the probability \( P \) is rather small: (i) negatively large threshold \( D \), (ii) small expiration time \( T \), and (iii) both. The statistics of Value-at-Risk, denoted by \( \text{VaR}_\alpha \), is the \((1 - \alpha)\times100\) percentile of log returns, where significance level \( 0 \leq \alpha \leq 1 \). Hence, the exact calculation of \( \text{VaR}_\alpha \) ends up solving a nonlinear equation
\[ 1 - \alpha = P(0, x, y; \text{VaR}_\alpha). \]
(4)
CVaR is simply a conditional expectation given that losses are greater than or equal to the \( \text{VaR}_\alpha \). Due to the complexity of stochastic volatility models, there is no closed-form solution in general for either the pricing or hedging value defined in (2) or the extremal event probability \( P \) defined in (3). Thus, computational challenges arise when one needs to obtain \( \text{VaR}_\alpha \) by solving Equation (4).

The tradeoff between estimation accuracy and computing speed is one of the major concerns. Fouque et al. (2000) derive explicit formula to approximate the value function \( P \) defined in (2) or (3) by means of perturbation asymptotics. This method is favored by its computing speed but at the cost of accuracy. Other numerical methods aim at estimation accuracy, including Monte Carlo method (Glasserman 2003) and numerical partial differential equations (Lipton 2001). We value the exact calculation of VaR and CVaR, and choose Monte Carlo method as our computational tool because this method possesses at least the following advantages: (i) an unbiased estimator for extremal event probability, and (ii) a straightforward generalization to calculate CVaR. However, Monte Carlo method is known for its deficiency of slow convergence. Variance reduction technique is a must when high precision is demanded. Given that importance sampling can increase the occurrence of extremal events, it is very useful for small probability estimation (Bucklew 2004; Han 2010). Section 4 is devoted to the study of importance sampling estimators for computing the extremal event probability defined in (3). Before proceeding to this computation, we estimate the parameters within stochastic volatility model (1) by means of Fourier transform method.
3 Volatility Estimation: Fourier Transform Method

Fourier transform method (Malliavin and Mancino 2002, 2009) is a nonparametric method to estimate multivariate volatility process. Its main idea is to reconstruct volatility as time series in terms of sine and cosine basis under the following continuous semi-martingale assumption. Let \( u_t \) be the log-price of an underlying asset \( S \) at time \( t \), i.e. \( u_t = \ln S_t \), and follow a diffusion process

\[
du_t = \mu_t dt + \sigma_t dW_t,
\]

where \( \mu_t \) is the instantaneous growth rate and \( W_t \) is a one-dimensional standard Brownian motion. One can estimate the time series volatility \( \sigma_t \) with the following steps.

- **Step 1:** Compute the Fourier coefficients of the underlying \( u_t \) as follows:

\[
a_0(du) = \frac{1}{2\pi} \int_0^{2\pi} du_t,
\]

\[
a_k(du) = \frac{1}{\pi} \int_0^{2\pi} \cos(kt) du_t,
\]

\[
b_k(du) = \frac{1}{\pi} \int_0^{2\pi} \sin(kt) du_t,
\]

for any \( k \geq 1 \), so that \( u(t) = a_0 + \sum_{k=1}^{\infty} \left[ -\frac{b_k(du)}{k} \cos(kt) + \frac{a_k(du)}{k} \sin(kt) \right] \). Note that the original time interval \([0,T]\) can always be rescaled to \([0,2\pi]\) as shown in above integrals.

- **Step 2:** Compute the Fourier coefficients of variance \( \sigma_t^2 \) as follows:

\[
a_k(\sigma^2) = \lim_{N \to \infty} \frac{\pi}{2N+1} \sum_{s=N}^{N-k} \left[ a_s^*(du) a_{s+k}^*(du) + b_s^*(du) b_{s+k}^*(du) \right],
\]

\[
b_k(\sigma^2) = \lim_{N \to \infty} \frac{\pi}{2N+1} \sum_{s=N}^{N-k} \left[ a_s^*(du) b_{s+k}^*(du) - b_s^*(du) a_{s+k}^*(du) \right],
\]

for \( k \geq 0 \), in which \( a_s^*(du) \) and \( b_s^*(du) \) are defined by

\[
a_s^*(du) = \begin{cases} a_s(du), & \text{if } s > 0 \\ 0, & \text{if } s = 0 \end{cases} \quad \text{and} \quad b_s^*(du) = \begin{cases} b_s(du), & \text{if } s > 0 \\ 0, & \text{if } s = 0 \end{cases}
\]

\[
a_s(du), \quad \text{if } s < 0 \quad \text{and} \quad b_s(du), \quad \text{if } s < 0
\]

- **Step 3:** Reconstruct the time series of variance \( \sigma_t^2 \) by

\[
\sigma_t^2 = \lim_{N \to \infty} \sum_{k=0}^{N} \phi(\delta k) [a_k(\sigma^2) \cos(kt) + b_k(\sigma^2) \sin(kt)],
\]
where $\varphi(x) = \frac{\sin^2(x)}{x^2}$ is a smooth function with the initial condition $\varphi(0) = 1$ and $\delta$ is a smooth parameter typically specified as $\delta = \frac{1}{50}$ (Reno 2008).

From Equations (6)-(8), it is observed that the integration error of Fourier coefficients is adversely proportional to data frequency. This Fourier transform method is easy to implement because, as shown in (9) and (10), Fourier coefficients of the variance time series can be approximated by a finite sum of multiplications of $a^*$ and $b^*$. This integration method can accordingly avoid drawbacks inherited from those traditional methods based on the differentiation of quadratic variation.

3.1 A Price Correction Scheme: Bias Reduction

It is documented that this Fourier transform method incurs a “boundary effect.” Reno (2008) notes that Fourier algorithm provides inaccurate estimate for volatility time series near the boundary of simulated data. He suggests that all the time series of estimated volatility near the first and last 1% should be discarded for the purpose of better estimation. To remedy this boundary deficit, we propose a price correction scheme as follows. Recall that $u_t$ defined in (5) is the natural logarithm of asset price. Based on the Euler discretization, the increment of log-price $u_t$ can be approximated by $\sigma_t \sqrt{\Delta_t} \varepsilon_t$, where $\Delta_t$ denotes a small discretized time interval and $\varepsilon_t$ denotes a sequence of i.i.d. standard normal random variables. This approximation is derived from neglecting the drift term of small order $\Delta_t$ and using the increment distribution of Brownian motion $\Delta W_t = \sqrt{\Delta_t} \varepsilon_t$. Given a set of discrete observations of log returns, let $\hat{\sigma}_t$ denote the volatility time series estimated from the original Fourier transform method. Our price correction scheme consists of a linear transformation on the natural logarithm of estimated variance process $\hat{\sigma}_t^2$ in order to guarantee the positiveness of estimated volatility. That is, we transform $\hat{Y}_t = 2 \ln \hat{\sigma}_t$ to $a + b \hat{Y}_t$ so that the corrected volatility $\sigma_t = \exp\left(\left(a + b \hat{Y}_t\right)/2\right) > 0$ satisfies $\Delta u_t \approx \exp\left(\left(a + b \hat{Y}_t\right)/2\right) \sqrt{\Delta_t} \varepsilon_t$, where $\Delta u_t = u_{t+1} - u_t$, and $a$ and $b$ denote the correction variables. This linear transformation on $\hat{Y}_t$ can be understood as the first order approximation to a possible nonlinear transformation on estimated volatility $\hat{\sigma}_t$. Then we can use the maximum likelihood method to regress out correction variables via the relationship between logarithm of squared standardized return $\Delta u_t / \sqrt{\Delta_t}$ and the driving volatility process $a + b \hat{Y}_t$:

$$\ln \left(\frac{\Delta u_t}{\sqrt{\Delta_t}}\right)^2 = a + b \hat{Y}_t + \ln \varepsilon_t^2. \quad (12)$$

To empirically test our price correction scheme, we set stochastic volatility model parameters of
Ornstein-Uhlenbeck type defined in (1) and (13) as follows: \( \mu = 0.01, S_0 = 50, \quad Y_0 = -2, m = -2, \quad \alpha = 5, \quad \beta = 1, \) and \( \rho = 0 \) with the discretization length \( \Delta_t = 1/5000, \) so as to generate volatility series \( \sigma_t = \exp(Y_t/2) \) and asset price series \( S_t. \) Based on (i) the original Fourier transform method and (ii) the one with our price correction scheme, we estimate two volatility time series, and compare them with the actual volatility series. Two criteria are used for performance comparison: Mean squared errors (MSE) and Maximum absolute errors (MAE). Comparison results are shown below:

1. Mean squared error: 0.0324 (Fourier method), 0.0025 (Corrected Fourier method)
2. Maximum absolute error: 0.3504 (Fourier method), 0.1563 (Corrected Fourier method).

Noticeably, our price correction scheme shown in (12) is easy to implement because it is based on a simple maximum likelihood estimation. It reduces effectively both error criteria at least by half in this simulated example. Similar performance results can be observed in other intensive simulations. Thus, it is worthwhile to include this price correction scheme in Fourier transform method to remedy the “boundary effect.” Up to this stage, the corrected Fourier transform method remains model-free under the semi-martingale framework (5) because no parameters within the stochastic volatility model appear on the regression equation (12).

3.2 Stochastic Volatility Model Estimation

Given that the volatility time series is estimated by corrected Fourier method, we proceed to estimate stochastic volatility model parameters. Assuming that the driving volatility process \( Y_t \) is governed by the Ornstein-Uhlenbeck process, i.e.

\[
dY_t = \alpha (m - Y_t) dt + \beta dW_t, \tag{13}
\]

we use the corrected estimation \( a + b \hat{Y}_t \) obtained from (12) to further estimate model parameters \( (\alpha, \beta, m) \) of \( Y_t \) by means of the maximum likelihood method. For a given set of observations \( Y_1, ..., Y_N, \) the likelihood function is

\[
L(\alpha, \beta, m) = \prod_{t=1}^{N} \frac{1}{\sqrt{2\pi \beta^2 \Delta_t}} \exp \left\{ \frac{-1}{2 \beta^2 \Delta_t} \left[ Y_{t+1} - (\alpha m \Delta_t + (1 - \alpha \Delta_t) Y_t) \right]^2 \right\}, \tag{14}
\]

where \( \Delta_t \) denotes the length of discretized time interval. This likelihood function is obtained by discretizing the stochastic differential equation (13). Taking the natural logarithm and ignoring the constant term, the log-likelihood becomes

\[
\ln L(\alpha, \beta, m) \propto N \ln \beta + \frac{1}{2 \beta^2 \Delta_t} \sum_{t=1}^{N-1} \left[ Y_{t+1} - (\alpha m \Delta_t + (1 - \alpha \Delta_t) Y_t) \right]^2.
\]

By maximizing the right hand side over the parameters \( (\alpha, \beta, m) \), we obtain the following maximum likelihood estimators
\[
\hat{\alpha} = \frac{1}{\Delta t} \left[ 1 - \frac{\left( \sum_{i=2}^{N} Y_i \right) \left( \sum_{i=1}^{N-1} Y_i \right) - (N-1) \left( \sum_{i=1}^{N-1} Y_i Y_{i+1} \right)}{\left( \sum_{i=1}^{N-1} Y_i \right)^2 - (N-1) \left( \sum_{i=1}^{N-1} Y_i^2 \right)} \right], \tag{15}
\]
\[
\hat{\beta} = \frac{1}{N \Delta t} \sum_{i=1}^{N-1} \left[ Y_{i+1} - \left( \alpha m \Delta t + (1 - \alpha \Delta t) Y_i \right) \right]^2, \tag{16}
\]
\[
\hat{m} = -\frac{1}{\hat{\alpha} \Delta t} \left[ \frac{\left( \sum_{i=2}^{N} Y_i \right) \left( \sum_{i=1}^{N-1} Y_i^2 \right) - \left( \sum_{i=1}^{N-1} Y_i \right) \left( \sum_{i=1}^{N-1} Y_i Y_{i+1} \right)}{\left( \sum_{i=1}^{N-1} Y_i \right)^2 - (N-1) \left( \sum_{i=1}^{N-1} Y_i^2 \right)} \right]. \tag{17}
\]
These are estimators of mean-reverting rate, volatility of volatility, and long-run mean, respectively.

4 Importance Sampling: Variance Reduction

When dealing with sparse observations in the tails, the basic Monte Carlo simulation is not favored for its undesirable properties, e.g. large relative error and data clustering around the center, etc. Importance sampling is one of the major methods of variance reduction to improve the convergence of basic Monte Carlo method. The fundamental idea behind importance sampling is to relocate the original density function to the area of interest with proper weights. The relocated density typically incurs more occurrence of rare events so that an accurate estimate for a small probability can be achieved. This technique is very helpful in rare event simulation. See Bucklew (2004) for discussions on importance sampling and extremal event probability estimation.

There are two major categories of the studies (Lemieux 2009) to investigate methods of importance sampling and their efficiency. The first category aims to reduce the variance of an importance sampling estimator as much as possible. This approach often ends up solving a fully nonlinear optimization problem, possibly in high dimension, or solving a simplified optimization problem derived from some approximation techniques. The second category of importance sampling methods emphasizes on minimizing the variance rate of an importance sampling estimator. The notion of variance rate is defined as the difference between the decay rate of the second moment and the decay rate of the square of the first moment. It has become a measure of efficiency for importance sampling estimators. When zero variance rate (note: not variance itself) is achieved, the corresponding importance sampling estimator is known as asymptotically optimal or efficient. This approach has been successfully applied to problems of rare event simulation.

Our importance sampling estimators emerge from the second category. This category offers noticeable advantages, as compared with the first category, including easier implementation, reduced computational cost, and more analytical tractability. The first two advantages are manifest in itself, as opposed to solving for high-dimensional nonlinear optimization problems in the first category. The third advantage helps link our proposed algorithm for extremal event simulations with large deviation theory (Bucklew 2004). This theory provides sharp estimates for the decay rate of small probabilities.
We proceed to introduce our proposed algorithm of importance sampling with respect to the Black–Scholes model and stochastic volatility models. According to the ergodic property of the averaged variance process, the constant volatility of the Black-Scholes model can be viewed as a limit of some stochastic volatility model. See Fouque et al. (2000) for details. Namely, one can treat a stochastic volatility model as a perturbation around the Black-Scholes model. In a way, to investigate an importance sampling estimator for stochastic volatility model, it is natural to first study the Black-Scholes model. Based on the large deviation principle of normal random variables, an efficient importance sampling and its variance analysis for the Black-Scholes model is established in Section 4.1. Under a stochastic volatility environment, we first carry out its limiting volatility, or called effective volatility, then apply the aforementioned importance sampling estimator for the Black-Scholes model to stochastic volatility models. This is detailed in Section 4.2.

4.1 Black-Scholes Model

Since the Black-Scholes model assumes that the risky asset price follows a geometric Brownian motion

\[ dS_t = \mu S_t dt + \sigma S_t dW_t, \]

its log return \( \ln(S_T/S_0) \) is normally distributed for \( T > 0 \). As a result, the extremal event probability with the threshold \( D \), denoted by \( P(0,S_0) \), admits the closed-form solution

\[
P(0,S_0) = E\left\{ \mathbf{1}\left( \ln(S_T/S_0) \leq D \right) \right\} S_0 \right| \\
= N\left( \frac{D - (\mu - \sigma^2/2)T}{\sigma \sqrt{T}} \right),
\]

where \( N(\cdot) \) denotes the cumulative normal integral function. We remark that in the case of VaR estimation, \( D \) is equal to \( \text{VaR}_\alpha \) so that \( E\left\{ \mathbf{1}\left( S_T \leq S_0 \exp(D) \right) \right\} = (1-\alpha) \times 100\% \).

A basic Monte Carlo method provides an unbiased estimator for the extremal event probability \( P(0,S_0) \) defined in (18) by the sample mean of extremal event indicators

\[
P(0,S_0) \approx \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\left( \ln(S^{(i)}_T/S_0) \leq D \right),
\]

where \( N \) is the total number of i.i.d. random samples and \( S^{(i)}_T \) denotes the \( i \)-th simulated asset price at time \( T \).

Next we investigate an efficient importance sampling estimator to estimate \( P(0,S_0) \). By Girsanov theorem (Oksendal 1998), one can construct an equivalent probability measure \( \bar{P} \) defined by the Radon-Nikodym derivative

\[
\frac{d\bar{P}}{dP} = Q_T(h) = \exp\left( \int_0^T h(s,S_s) d\bar{W}_s - \frac{1}{2} \int_0^T h(s,S_s)^2 ds \right), \text{ where } \bar{W}_t = W_t + \int_0^t h(s,S_s) ds
\]

is a...
Brownian motion under $\mathbb{P}$ provided that the process $h(s, S_t)$ satisfies, say, Novikov’s condition to ensure certain integrability of the function $h$ such that $Q_t$ is a martingale for $0 \leq t \leq T$.

The proposed importance sampling scheme is determined by a constant drift change $h$ in order to satisfy the intuition of “the expected asset value $S_T$ under the new probability measure is equal to its threshold $S_0 \exp(D)$,” i.e.

$$\tilde{E}\{S_T|F_0\} = S_0 \exp(D).$$  \hspace{1cm} (20)

This intuition can be rigorously verified based on the construction of exponential change of measure (Glasserman 2003; Han 2010). Hence, the extremal event $|\ln(S_T/S_0)\leq D\}$, when $D$ is negatively large and/or $T$ is small, is no longer rare under the new probability measure and the accuracy of Monte Carlo simulation can be improved significantly. Using the log-normal density of $S_T$, the criterion (20) results in a unique drift change

$$h = \frac{\mu - D}{\sigma T}. \hspace{1cm} (21)$$

Therefore, under the new probability measure $\mathbb{P}$ defined from the Radon-Nikodym derivative

$$Q_T(h) = \exp\left(h\bar{W}_T - \frac{h^2T}{2}\right), \hspace{1cm} (22)$$

the extremal probability defined in (18) can be re-expressed as

$$P(0, S_0) = \tilde{E}\{\mathbf{1}\{\ln(S_T/S_0)\leq D\} Q_T(h)|S_0\}, \hspace{1cm} (23)$$

where the underlying risky-asset process is governed by $dS_t = (\mu - \sigma h)S_t dt + \sigma S_t d\bar{W}_t$ due to this change of measure. The unbiased importance sampling estimator of $P(0, S_0)$ is

$$\frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{\ln(S_T^{(i)}/S_0)\leq D\} Q_T^{(i)}(h). \hspace{1cm} (24)$$

The following theorem confirms that our proposed importance sampling estimator (24) is asymptotically optimal or efficient. We show that its variance rate approaches zero when extremal events occur. That is, the decay rate of the second moment of $\mathbf{1}\{\ln(S_T/S_0)\leq D\} Q_T(h)$ is twice of the decay rate of its first moment under some scaling scenarios. Thus, we can present the following theorem for our importance sampling method.

11
Theorem 1 Under the Black-Scholes model, the variance rate of the proposed importance sampling scheme defined in (23) approaches zero in cases of (i) very short maturity, (ii) default threshold is negatively large, or (iii) both. That is, the proposed importance sampling estimator (24) is asymptotically optimal or efficient under some scaling scenarios in time and space.

The complete proof set, involving decay rate estimation of the first and second moments of importance sampling estimators under a spatial scale and/or a time scale, is exhibited in the Appendix A. Our discussion on importance sampling not only serves as an alternative computation for the closed-form solution (18), but also lays a foundation to treat similar problems under stochastic volatility models.

4.2 Stochastic Volatility Model

In general, there is no closed-form solution for the evaluation problem defined in (2) under the stochastic volatility model (1). Monte Carlo simulation is a generic approach to solve for this problem. In last two decades, many attentions focus on variance reduction techniques under stochastic volatility models. Willard (1996) develops a conditional Monte Carlo scheme by conditioning on the driving volatility process. Fournie et al. (1997) and Fouque and Han (2004) apply regular and/or singular perturbation techniques to develop methods of importance sampling, control variate, or estimators combined with these two methods. Heath and Platen (2002) use an option price approximation with deterministic volatility to construct a control variate method. Fouque and Han (2007) generalize this approach to option price approximation with random volatility under multi-factor stochastic volatility models, and provide a variance analysis given two well-separated time scales. Han and Lai (2010) develop generalized control variate methods combined with Quasi Monte Carlo for enhanced reduction of variance. These control variate methods and importance sampling are more computationally intensive because a priori approximation to the evaluation problem (2) is used in these algorithms.

Here we propose an alternative importance sampling estimator to estimate the extremal event probability under stochastic volatility models. The pro is that no prior knowledge about the unknown quantity defined in (2) is required, as opposed to perturbation approximation discussed previously, so the computational cost, e.g. the execution time of this new estimator, is significantly reduced. The con is that the sample variance of this new estimator may be greater than those obtained from perturbation methods. However, in our simulation and empirical studies we find that the proposed estimator can indeed produce efficient and accurate estimation for the extremal event probability.

The assumption of our proposed importance sampling estimators under stochastic volatility models is that the following ergodic property of the average of the variance process

$$\frac{1}{T} \int_0^T f(Y_t^ε)^2 dt \overset{a.s.}{\to} \bar{σ}^2, \text{for } ε \to 0$$

(25)

holds, where $ε$ denotes a small time scale and $Y_t^ε$ denotes a fast mean-reverting process. The effective volatility $\bar{σ}$ is a constant defined by the square root of the expectation of the variance function $f(\cdot)^2$ with
respect to the limiting distribution of $Y_t^\varepsilon$, namely, $\bar{\sigma}^2 = \int f^2(y) \, d\Phi(y)$, where $\Phi(y)$ denotes the invariant distribution of the fast varying process $Y_t^\varepsilon$. Ornstein-Uhlenbeck process is a typical example of the stochastic volatility model. Under the fast mean-reverting assumption, coefficient functions of $Y_t$ defined in (1) are chosen as $c(y) = \frac{1}{\varepsilon} (m - y)$ and $g(y) = \sqrt{2\nu / \varepsilon}$, so that the invariant distribution $\Phi$ is simply a Gaussian with mean $m$ and variance $\nu$. These results are thoroughly discussed in Fouque et al. (2000).

The limiting result (25) suggests a change of probability measure as follows. By substituting $\bar{\sigma}$ into $\sigma$ shown in (21), one can define $Q_T(h(\bar{\sigma})) = \exp \left( h(\bar{\sigma}) \bar{W}^{(0)}_t - \frac{1}{2} h^2(\bar{\sigma}) T \right)$ as a Radon-Nykodym derivative so that $\bar{W}^{(0)}_t = W^{(0)}_t + h(\bar{\sigma}) t$ is a Brownian motion under the new measure denoted by $\bar{\mathcal{P}}$. Therefore, the extremal event probability defined in (3) can be re-expressed as

$$P(0, S_0, Y_0) = \bar{E}\{\ln(S_T / S_0) \leq D \mid Q_T(h(\bar{\sigma})) \mid S_0, Y_0\},$$

where the underlying risky-asset process is governed by

$$dS_t = (\mu - \sigma h(\bar{\sigma})) S_t \, dt + \sigma S_t \, d\bar{W}^{(0)}_t$$

and the dynamics of $Y_t$ is changed accordingly. The unbiased importance sampling estimator for $P(0, S_0, Y_0)$ becomes

$$\frac{1}{N} \sum_{i=1}^{N} I(\ln(S^{(i)}_T / S_0) \leq D) Q^{(i)}_T(h(\sigma)).$$

(27)

As a matter of fact, the proposed importance sampling can only affect the prominent Brownian motion $W^{(0)}_t$ defined in (1) while the other independent Brownian motion $W^{(1)}_t$ appearing in the volatility process $Y_t$ is not affected by this measure change.

This partial change of measure takes the advantage of some averaging effect of randomness, as seen in (25). A similar idea emerges from a recent study Han (2010) of importance sampling for multivariate Student-t variate, in which a normalized Chi square random variable can be considered as part of the random volatility and it converges to one when the degree of freedom becomes large by the law of large numbers. It is known that the distribution of Student-t variate converges to a Gaussian (Finner et al. 2008). This limiting effect leads to a partial change of measure. That is, when one designs an importance sampling scheme for estimating the extremal event probability of Student-t, using change of measure solely for Gaussian rather than Student-t itself can actually be a good choice.
4.3 Conditional VaR

Artzner et al. (1999) provide some criteria for qualifying a coherent risk measure. While VaR fails, CVaR, also known as expected shortfall, is qualified as a coherent risk measure. CVaR is defined as a conditional expectation $E\{X | X < c\}$, where $X$ denotes a loss variable and $c = VaR_a$ satisfies $E\{I(X \leq VaR_a)\} = (1-\alpha) \times 100\%$. CVaR is the conditional expectation of the tail.

The basic Monte Carlo algorithm to calculate CVaR, i.e. $E\{X | X < c\}$ is as follows:

$$n_c = \sum_{i=1}^{N} I(X^{(i)} < c),$$

where $N$ is the total number of simulations.

$$E\{X | X < c\} \approx \frac{1}{n_c} \sum_{i=1}^{n_c} X^{(i)}, \text{ for each } X^{(i)} < c.$$  \hspace{1cm} (28)

By choosing a likelihood function $Q = \frac{dP}{d\tilde{P}}$, a new probability measure $\tilde{P}$ is defined. One can derive the following importance sampling estimator for CVaR:

$$E\{X | X < c\} = \frac{E\{\lambda I(X < c)\}}{E\{I(X < c)\}},$$

$$= \frac{\tilde{E}\{\lambda I(X < c)Q\}}{\tilde{E}\{I(X < c)Q\}},$$

$$\approx \frac{1}{N} \sum_{i=1}^{N} X^{(i)} I(X^{(i)} < c)Q(X^{(i)})$$

$$= \frac{1}{N} \sum_{i=1}^{N} I(X^{(i)} < c)Q(X^{(i)})$$

$$= \sum_{i=1}^{n_c} X^{(i)} q_i, \text{ for each } X^{(i)} < c,$$ \hspace{1cm} (29)

where $q_i = Q(X^{(i)}) / \sum_{i=1}^{n_c} Q(X^{(i)})$. We have used the definition of CVaR (first line in (29)), the same change of probability measure (second line), the same Monte Carlo simulation (third line), and straightforward calculation. Note that under this importance sampling algorithm, CVaR is approximated by the sum of a collection of random sample $X^{(i)}$, $i=1,\cdots,n_c$, multiplied by its corresponding weight $q_i$. This is different from the basic Monte Carlo estimator defined in (28), in which the weight associated with each random sample is equal to $1/n_c$ uniformly. We remark that approximation of the standard error (s.e.) in this
Non-equally weighted case is \( \text{s.e.} \approx \sqrt{\frac{1}{n_c} \sum_{i=1}^{n_c} (X^{(i)} - \hat{m})^2 q_i} \), for each \( X^{(i)} < c \), where \( \hat{m} \) denotes the sample mean. Next we show one asymptotic property of CVaR under the standard normal assumption.

**Lemma 2** When \( X \) is a standard normal, \( \lim_{c \to -\infty} \frac{E\{X \mid X < c\}}{c} = 1. \) This implies that when significance level \( \alpha \) approaches one, i.e. \( c = VaR_\alpha \) approaches negative infinity, \( CVaR = E\{X \mid X < c\} \) is asymptotically equal to its \( VaR_\alpha \).

The proof can be obtained directly by using the exact calculation

\[
E\{X \mid X < c\} = \frac{-e^{-c^2/2}}{\sqrt{2\pi} N(c)}
\] (30)

and the approximation \( N(c) \approx \frac{-1}{\sqrt{2\pi} c} e^{-c^2/2} \) for negatively large \( c \). This lemma provides an alert that using a normal approximation can possibly underestimate CVaR particularly when fat- (or fatter-) tailed distribution is considered in the model. Table 2 in following Section 4.4 summarizes the numerical comparisons about the discrepancy to estimate CVaR by using a normal approximation and the proposed importance sampling under a stochastic volatility model.

### 4.4 Numerical Examples

Two sets of numerical experiments are conducted to demonstrate efficiency of proposed importance sampling algorithms. The first set takes the Black-Scholes model as a benchmark. Closed-form solutions for extremal event probabilities (associated with \( VaR_{95\%} \) and \( VaR_{99\%} \)) and their CVaRs are used to compare numerics estimated from the basic Monte Carlo and importance sampling. The second set concentrates on CVaR estimation under a stochastic volatility model with various values of correlation. CVaRs are calculated from (i) a normal approximation and (ii) importance sampling.

In Table 1, we compare numerical results, obtained from the exact solution, the basic Monte Carlo, and the importance sampling algorithm, for estimating two extremal event probabilities and CVaRs given two loss threshold level \( D = -0.0313 \) and \( D = -0.0441 \), respectively, under the Black-Scholes model assumption \( dS_t = \mu S_t dt + \sigma S_t dW_t \). The model parameters are set \( \mu = 0, \sigma = 0.3 \) and \( T = 1/252 \) (one day), and the total number of simulation is 1,000,000. Given these setup, we calculate estimation of extremal event probability and its CVaR with two different loss thresholds \( (D) \). In the first column, two loss thresholds are exactly \( VaR_{95\%} \) and \( VaR_{99\%} \), so that extremal event probabilities are 0.05 and 0.01, respectively. Standard errors obtained from importance sampling are all smaller than those from basic Monte Carlo. Variance reduction ratios for extremal event probability and CVaR estimation are ranged from 4 to 12 in the case of
$D = -0.0313$ and from 36 to 60 in the case of $D = -0.0441$. Importance sampling gives significant performance when the extremal event probability is small, say 0.01, in this simulation study. Note that each estimated CVaR is close to the loss threshold $D$. This numerical result coincides with the approximation predicted by Lemma 2.

Table 2 demonstrates how stochastic volatility models contribute to the expected shortfall CVaR relative to a normality assumption given the same level of VaR. Parameters of the stochastic volatility model defined in (1) and (13) with $\sigma_t = \exp(Y_t/2)$ are chosen as $\mu = 0$, $m = -5$, $\alpha = 5$, $\beta = 1$, $S_0 = 50$, $Y_0 = -3$, $T = 1/252$ (one day), and the confidence level of VaR is $\alpha = 0.99$. The correlation between Brownian motions is varied and listed in Column 1, and $VaR_{99\%}$ are estimated and listed in Column 2. The total number of simulation for basic Monte Carlo defined in (28) and the importance sampling defined in (29) is uniformly set as 10,000. CVaR defined as $E\left[\ln\left(\frac{S_T}{S_0}\right)\ln\left(\frac{S_T}{S_0}\right) < VaR_{99\%}\right]$ is calculated by two methods: a normal approximation and importance sampling, given that $VaR_{99\%}$ has been estimated separately from another importance sampling simulation. The normal approximation, denoted by N. Approx. in Column 3, is derived based on the Black-Scholes model specified as (18) with the spot volatility $\sigma = \exp(Y_0/2)$, where $Y_0$ is the initial value of driving volatility process. That is, we approximate CVaR by assuming that $-\ln(S_T/S_0)$ is normally distributed with mean $(\exp(Y_0)/2 - \mu)T$ and variance $\exp(Y_0)T$. Column 4 reports CVaR estimated from importance sampling with standard errors. These standard errors are all smaller relative to CVaR estimates. This contrast shows the accuracy (efficiency) of the proposed importance sampling. It is worth noting that relative errors between the normal approximation (Column 3) and importance sampling (Column 4) range from 10.03% to 12.62%. These discrepancies are significant and are consequences of the fatted tail brought by the stochastic volatility model.

In summary, Table 1 confirms the outperformance of our enhanced importance sampling method under the Black-Scholes model. Table 2 exhibits more accurate CVaR estimates by this method than normal approximation under stochastic volatility models. We can conclude that our proposed importance sampling method can indeed provide accurate estimation.

Note that under the configuration of our PC (Intel CPU Core 2 Duo 2.4GHz), it is rather time-consuming to solve for VaR under stochastic volatility models without importance sampling. Under Matlab environment, we use the basic Monte Carlo estimator to approximate the extremal event probability defined in (3), then use a nonlinear Matlab solver, say fzero.m, to solve for $VaR_{99\%}$ defined in (4). Even if the number of simulation increases to 250,000 and the execution time exceeds three minutes, we are still unable to solve for a single $VaR_{99\%}$ estimate. This indicates that VaR estimation is a challenging task for Monte Carlo simulation under stochastic volatility models. In contrast, our proposed importance sampling algorithm takes only few seconds to get a $VaR_{99\%}$ estimate. Further, without a variance reduction, using the basic Monte Carlo method to estimate VaR is expected to consume tremendous computing resources. Our importance sampling algorithm provides even more significant advantage of saving computing efforts and execution.
5 Backtesting for VaR Estimation

Backtesting helps avoid model misspecification and differentiate the model performance from a faulty model under special conditions. In effect, backtesting can balance Type I against Type II statistical errors in VaR estimation. There are two major criteria for backtesting: unconditional rate of exceedances (UC) and independence of the exceedances (IND). It is expected that the significance level represents the maximum probability of observations exceeding VaR estimates if the model is correctly calibrated.

Under the null hypothesis that the significance level is the true probability of exceedances occurring, the test statistics are a log-likelihood ratio specified as:

\[
LR_{UC} = -2 \ln \left( (1-p)^{T-N} \ p^N \right) + 2 \ln \left\{ \left( 1 - \frac{N}{T} \right)^{T-N} \left( \frac{N}{T} \right)^N \right\} \sim \chi^2(1),
\]

where \( T \) is total number of days and \( N \) is the number of exceedances. This asymptotically follows a Chi-square distribution with one degree of freedom (Kupiec 1995).

For the independence test of the exceedances, the first job is to set up a series which indicates if the daily VaR estimate is exceeded or not. If the VaR estimate is not exceeded by the actual loss, the exceedance indicator is set at 0, or 1 otherwise. The next job is to observe the switches of exceedances. Table 3 shows how to construct a table of conditional exceedances. The log-likelihood test statistics are specified as:

\[
LR_{IND} = -2 \ln \left( 1 + \pi(T_{00}+T_{10}) \ \pi(T_{01}+T_{11}) \right) + 2 \ln \left( (1 - \pi)_{00} \ \pi_{01} (1 - \pi)_{01} \ \pi_{11} \right) \sim \chi^2(1),
\]

where \( T_j \) denotes the number of days in which state \( j \) occurred in one day while it was state \( i \) the previous day. Moreover, \( \pi_j \) represents the probability of observing an exceedance conditional on state \( i \) the previous day. It asymptotically follows a Chi-square distribution with one degree of freedom. The first term is specified under the hypothesis that the exceedances are independent across the sample, or

\[
T_1 = T_0 = T_{11} = \frac{(T_{00} + T_{11})}{T}.
\]

The second term is the maximized likelihood for the observed data. This test helps confirm if the exceedances are serially correlated, i.e. to examine whether the model makes systematic errors in the VaR estimates.

The conditional coverage (CC) test is designed to simultaneously test if the VaR violations are independent and the average number of exceedances is correct. The test statistics for conditional coverage are actually the sum of the test statistics of unconditional coverage and independence, i.e.
These three types of backtesting- unconditional coverage, independence, and conditional coverage- are popularly practiced to help validate VaR estimation performance (Christoffersen 1998).

5.1 Empirical Analysis

VaR and CVaR, two of the most widely used risk measures, are used as the criteria for performance comparison. The competing methods include historical simulation, RiskMetrics\(^4\), and GARCH\((1,1)\)\(^5\). The first one is acknowledged as popular and easy to implement and free of assumption. The latter two are known for being succinct and robust in capturing volatility process. All of them are commonly accepted as benchmark models for VaR estimation.

Two datasets are used for empirical examination. The first dataset contains three exchange rate series against the US Dollar: Japanese Yen (JPY), Singapore Dollar (SGD), and Canadian Dollar (CAD). The data covers the period from January 5, 1998 to July 24, 2009, with 2890 daily observations. The dataset is collected from the Central Bank of the Republic of China (http://www.cbc.gov.tw/). The daily observations are taken as the natural logarithmic returns in two consecutive trading days and are denoted as

\[ X_t : X_t = \ln \left( \frac{r_t}{r_{t-1}} \right) , \text{ where } r_t \text{ is the daily exchange rate at date } t. \]

Descriptive statistics and time series plot of the six series (3 original series and 3 corresponding return series) are summarized in Table 4 and Figure 1, respectively. Figure 2 shows the estimates of the three major stochastic volatility model parameters in the respective return series by the proposed corrected Fourier transform scheme: \( \alpha \) (right), \( \beta \) (bottom left), and \( m \) (top-left) which represents mean-reverting rate, volatility of volatility, and long-run mean, respectively in the volatility process. We remark that the initial value of driving volatility \( Y_0 \) is determined by the corrected estimate \( a + b\hat{Y}_0 \), where \( \hat{Y}_0 \) is estimated from the original Fourier transform method. These three model parameters are designed to capture the volatility process, and the correction scheme offers bias reduction and relieves the “boundary effect” over traditional Fourier transform method. We find that estimated mean-reverting rates are consistent with the fast mean-reverting assumption postulated in (25). Noteworthy, those estimates of the three parameters in each series show significant spikes near the data period end. This indicates the active property of our correction scheme and its capacity

\[ LR_{CC} = LR_{UC} + LR_{IND}. \]

---

\(^4\) RiskMetrics is also called exponentially weighted moving average method and this method is designed to represent the finite-memory property. This method is specified to model the volatility process as:

\[ \sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2, \text{ where } r_{t-1} \text{ denotes return rate at time } (t-1). \]

The decaying factor (\( \lambda \)) in RiskMetrics model is set as 0.94 throughout this paper.

\(^5\) For return series \( r_t \), \( F_{t-1} = \mu_t + \sigma_t \epsilon_t, \epsilon_t \sim N(0,1) \), where \( F_{t-1} \) : information set, \( \mu_t \) : conditional mean, and \( \sigma_t^2 \) : conditional variance. GARCH\((1,1)\) model is specified as:

\[ \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta \sigma_{t-1}^2. \]

VaR estimation under GARCH\((1,1)\) can be found in Hull (2008).
to capture the dynamic variance along the major crisis since late 2007. The estimates of VaR and CVaR by the proposed importance sampling algorithm at 95% and 99% significance levels are plotted in Figure 3. In general, those VaR estimates can serve as a safety net for risk management purpose.

Three methods of backtesting (unconditional coverage, independence, and conditional coverage) are employed for performance examination and the outcomes are summarized in Table 5. The significance level for rejecting backtesting is set as 10%. The proposed improved procedure, including the corrected Fourier transform scheme and importance sampling scheme, dominates at 99% significance level. The overall evidence shows that this procedure overwhelmingly outperforms the aforementioned major traditional tools, with 2 exceptions at 95% significance level. The outperformance of this procedure is manifest at 99% extremal significance level. GARCH is evaluated as a competent competitor but its satisfactory performance is constrained at 95% significance level. The underperformance of historical simulation and RiskMetrics can be attributed to their gradual and rigid structure of adjustment to the volatility process.

The second dataset, downloaded from Yahoo! Finance website (http://finance.yahoo.com/), is composed of two series: daily observations of S&P 500 and its VIX (Chicago Board Options Exchange Volatility Index) data. The data coverage expands from January 3, 2005 to July 24, 2009, consisting of 1138 observations. VIX is an annualized volatility index of S&P 500. It is used to predict the market volatility level of the following 30 calendar days. With respect to the recent worldwide financial tsunami since 2007, the second dataset is selected to stress test the proposed procedure. The first 500 observations, which time period is before the major financial crisis, are used as a warm-up period for estimation. Both series are treated the same manner as that of the first dataset to obtain return series. Descriptive statistics and time series plot of both series are summarized in Table 6 and Figure 4, respectively. There is noticeable correlation between the two series (-0.67839 for original series and -0.71487 for log return series) and VIX is noticed as a major volatility index for the S&P 500 stock index. These properties are expected auxiliary in VaR estimation of S&P 500 series. Accordingly, VIX data functions for Rho estimation for our proposed procedure, rather than to constitute a bivariate portfolio.

We first use the S&P 500 and VIX data in the corresponding moving windows (fixed size of 500 daily data) to calculate the correlation (\( \rho \)) between them (Figure 5) and introduce the estimated \( \rho \) into the proposed correction scheme as for the foreign exchange rate series\(^6\). The parameter estimates (\( \alpha, \beta, m, \) and \( \rho \)) of this scheme for S&P 500 are demonstrated in Figure 6. Again, the rugged and spiky curves demonstrate the intense volatility along the estimation process, especially since the crisis period. The estimates of VaR and CVaR are plotted in Figure 7. The backtesting outcomes also give favorable conclusion on our proposed improved procedures performance (Table 7). The outperformance is especially significant at extremal 1% quantile level, which is stipulated by the Basel II Accord. The vibrant dynamics during the financial crisis are satisfactorily captured by our proposed method. Overall, our proposed procedure outperforms in VaR and CVaR estimation under stochastic volatility models as compared to traditional benchmark methods, i.e. historical simulation, RiskMetrics, and GARCH(1,1).

\(^6\) \( \rho \) is one of the major parameters in our proposed procedure. In contrast, we assume and set correlation equals zero, i.e. \( \rho=0 \), for the respective series in the first dataset which is designed for univariate VaR estimation.
6 Conclusion

VaR and CVaR estimation under stochastic volatility models is investigated. We tackle two major hurdles including (i) unstable parameter estimation subject to practical data constrain, and (ii) lack of efficient computational technique for accurately estimating extremal event probabilities. The first hurdle is overcome by a bias reduction procedure for Fourier transform method to estimate volatility via a price correction scheme, and the second hurdle is overcome by a variance reduction procedure for VaR and CVaR estimation via importance sampling.

Two datasets are selected for empirical examination; the first contains three exchange rate series and the second includes S&P500 index and its VIX. The long data period covers recent financial turmoil so as to stress test the VaR and CVaR estimation performance and examine their performance in capturing the stochastic volatility. In an empirical study, three essential types of backtesting are preceded for performance evaluation: unconditional coverage, independence, and conditional coverage. Backtesting outcomes show that our improved procedure under stochastic volatility models outperforms significantly at the 99% VaR estimation over classical benchmark methods: RiskMetrics, historical simulation, and the GARCH(1,1) model. This indicates practical contribution to the Basel II Accord implementation.
Appendix A: Proof of Efficient Importance Sampling Estimator

**Proof.** Firstly, we derive closed-form solutions for the first and second moments, denoted by $P_1$ and $P_2$ respectively, of any importance sampling scheme induced by a constant drift change $h$. Since $S_t = S_0 \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right)$ is log-normally distributed, we obtain the following closed-form solution for $P_1$:

$$P_1 = E \left\{ I(\frac{S_t}{S_0} \leq D) \right\} = N \left( \frac{D - (\mu - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}} \right).$$

From Equation (23), the second moment $P_2$ is defined by

$$P_2 = E \left\{ I(\frac{S_t}{S_0} \leq D) Q_\tau^2(h) \right\},$$

where the measure change is given by $Q_\tau(h) = \exp \left( h \tilde{W}_\tau - \frac{h^2 T}{2} \right)$, so that the asset price dynamics becomes

$$dS_t = (\mu - \sigma h)S_t dt + \sigma S_t d\tilde{W}_t$$

under the new probability measure $\tilde{P}$. Rewrite $P_2$ as

$$P_2 = e^{h^2 T} E \left\{ I(\frac{S_t}{S_0} \leq D) e^{2h \tilde{W}_\tau - \frac{12h^2 T}{2}} \right\} = e^{h^2 T} E \left\{ I(\frac{S_t}{S_0} \leq D) \right\},$$

where $\exp \left( 2h \tilde{W}_\tau - \frac{(2h)^2 T}{2} \right)$ is the Radon-Nykodym derivative to further change probability measure from $\tilde{P}$ to $\hat{P}$ such that $\tilde{W}_\tau := \tilde{W}_\tau - 2ht$ is a standard Brownian motion. Hence under $\hat{P}$, the dynamics of $S_t$ becomes $dS_t = (\mu + \sigma h)S_t dt + \sigma S_t d\tilde{W}_t$, so that we get the closed form for $P_2$:

$$P_2 = e^{h^2 T} N \left( \frac{D - (\mu + \sigma h - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}} \right).$$

In order to obtain zero variance rate, the key step is to choose a peculiar drift-change parameter $h$. 

21
According to (21), $h = \frac{\mu T - D}{\sigma T}$ so that the associated probability measure can incur more extremal events.

Secondly, we estimate decay rates of $P_1$ and $P_2$ under three scaling scenarios in time and space. When time scale $T$ is small, we set $T = \epsilon \ll 1$. It is easy to see that $\frac{D - (\mu - \sigma^2/2)T}{\sigma \sqrt{T}} \approx \frac{D}{\sigma \sqrt{\epsilon}}$ so that

$$P_1 \approx \frac{\sigma \sqrt{\epsilon}}{2\pi} e^{-\frac{D^2}{2\sigma^\epsilon}}$$

by using the normal approximation $N(-x) \approx \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}}$, where $x$ is positively large.

Because $h \approx \frac{D}{\sigma \epsilon}$ and $\frac{D - (\mu + \sigma h - \sigma^2/2)T}{\sigma \sqrt{T}} \approx \frac{2D}{\sigma \sqrt{\epsilon}}$, the approximation $P_2 \approx e^{\frac{D^2}{2\sigma^\epsilon}} N\left(\frac{2D}{\sigma \sqrt{\epsilon}}\right) \approx \frac{\sigma \sqrt{\epsilon}}{2\pi} e^{-\frac{2D^2}{2\sigma^\epsilon}} \approx e^{\frac{D^2}{2\sigma^\epsilon}} e^{\frac{-D^2}{\sigma^\epsilon}}$ is obtained. Therefore we get the following decay rates for the first two moments under a small time scale $T$:

$$\lim_{\epsilon \to 0} \epsilon \log P_1 = -\frac{D^2}{2\sigma^2},$$

$$\lim_{\epsilon \to 0} \epsilon \log P_2 = -\frac{D^2}{\sigma^2}.$$ 

These results show that the decay rate of the second moment is twice of the decay rate of the first moment, which implies $P_1^2 \approx P_2$ as $\epsilon$ goes to zero, so that an asymptotic zero variance rate for the importance sampling (24) is justified.

Similar results can be obtained under a small spatial scale, i.e. $D = -1/\sqrt{\epsilon}$ for $\epsilon \ll 1$. It is easy to check that

$$P_1 \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma T \epsilon}}$$

and

$$P_2 \approx e^{-\frac{1}{\sigma^2 T \epsilon}} N\left(\frac{-2}{\sigma \sqrt{T \epsilon}}\right) \approx e^{-\frac{1}{\sigma^2 T \epsilon}} \frac{1}{\sqrt{2\pi}} e^{\frac{2}{\sigma^2 T \epsilon}}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{\sigma^2 T \epsilon}}.$$ 

By inspection, $\lim_{\epsilon \to 0} \epsilon \log P_2 = 2 \lim_{\epsilon \to 0} \epsilon \log P_1 = -\frac{1}{\sigma^2 T}$ is obtained so that an asymptotic zero variance rate is
confirmed.

When maturity is short and default threshold is large, one can expect the increase of decay speed of these moments. Let \( D = -1/\sqrt{\varepsilon} \) and \( T = \varepsilon \) for \( \varepsilon << 1 \), then one can obtain the following decay rate estimates

\[
\lim_{\varepsilon \to 0} \varepsilon^2 \log P_2 = 2 \lim_{\varepsilon \to 0} \varepsilon^2 \log P_1 = -\frac{1}{\sigma^2}.
\]

Note that the scaling order is \( \varepsilon^2 \) in this scenario which is faster than \( \varepsilon \) in previous two scenarios.
Table 1: Estimation of Extremal Event Probability and Its CVaR with Two Different Loss Thresholds.

<table>
<thead>
<tr>
<th>Loss Threshold ($D$)</th>
<th>Extremal Event Probability</th>
<th>CVaR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact</td>
<td>BMC</td>
</tr>
<tr>
<td>-0.0313</td>
<td>0.0500</td>
<td>0.0510 (0.0002)</td>
</tr>
<tr>
<td>-0.0441</td>
<td>0.0100 (9.88E-0.5)</td>
<td>0.0099 (1.64E-0.5)</td>
</tr>
</tbody>
</table>

Remark:
1. Extremal event probability is defined as $P(0, S_0) = E[I(\ln(S_T/S_0) \leq D)]$ and its CVaR is defined as $E[\ln(S_T/S_0)\ln(S_T/S_0) \leq D]$. 
2. Exact, BMC, IS, and VR denote closed-form solution, Basic Monte Carlo method, importance sampling, and variance reduction ratio, respectively. VR is defined as $VR = \left(\frac{s.e. of BMC}{s.e. of IS}\right)^2$. 
3. Sample means and standard errors shown in parenthesis are reported in columns of BMC and IS.

Table 2: Comparisons of CVaR Approximation Based on Normal Approximation and Importance Sampling.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$c = VaR_{99%}$</th>
<th>CVaR</th>
<th>N. Approx.</th>
<th>IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>-0.0339</td>
<td>-0.0347</td>
<td>-0.0386 (7.5073E-05)</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>-0.0335</td>
<td>-0.0343</td>
<td>-0.0378 (7.3833E-05)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-0.0323</td>
<td>-0.0331</td>
<td>-0.0367 (7.2498E-05)</td>
<td></td>
</tr>
<tr>
<td>-0.4</td>
<td>-0.0317</td>
<td>-0.0325</td>
<td>-0.0366 (7.2739E-05)</td>
<td></td>
</tr>
<tr>
<td>-0.8</td>
<td>-0.0310</td>
<td>-0.0319</td>
<td>-0.0351 (6.9643E-05)</td>
<td></td>
</tr>
</tbody>
</table>

Remark:
1. Given five scenarios of correlation $\rho$ listed in Column 1 and estimated $VaR_{99\%}$ reported in Column 2, two sets of CVaRs are approximated. The first set of CVaR, reported in the Column 3, is an approximation obtained from the closed-form solution under a normality assumption. N. Approx. denotes the normal approximation.
2. The second set of CVaR is estimated by the proposed importance sampling estimator, denoted by IS. Sample means and their standard errors shown in parentheses are reported in the Column 4.

Table 3: Construction of Conditional Exceptions

<table>
<thead>
<tr>
<th>Current day</th>
<th>No exception</th>
<th>exception</th>
<th>unconditional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day before</td>
<td>$T_{00} = T_0 (1 - \pi_0)$</td>
<td>$T_{10} = T_1 (1 - \pi_1)$</td>
<td>$T (1 - \pi)$</td>
</tr>
<tr>
<td>no exception</td>
<td>$T_{0i} = T_0 (\pi_0)$</td>
<td>$T_{1i} = T_1 (\pi_1)$</td>
<td>$T (\pi)$</td>
</tr>
<tr>
<td>total</td>
<td>$T_0$</td>
<td>$T_1$</td>
<td>$T = T_0 + T_1$</td>
</tr>
</tbody>
</table>

Remark: $T_y$ denotes the number of days in which state j occurred in one day while it was state i the previous day. $\pi_i$ represents the probability of observing an exceedances conditional on state i the previous day.

Table 4: Descriptive Statistics of the three Foreign Exchange Rate Data

Panel 1: Original Daily Data

<table>
<thead>
<tr>
<th></th>
<th>JPY</th>
<th>SGD</th>
<th>CAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>87.915</td>
<td>1.34665</td>
<td>0.9218</td>
</tr>
<tr>
<td>1st Quantile</td>
<td>107.525</td>
<td>1.57576</td>
<td>1.1683</td>
</tr>
<tr>
<td>Mean</td>
<td>114.5035</td>
<td>1.653406</td>
<td>1.334586</td>
</tr>
<tr>
<td>Median</td>
<td>115.265</td>
<td>1.6875</td>
<td>1.357725</td>
</tr>
<tr>
<td>3rd Quantile</td>
<td>120.35</td>
<td>1.737975</td>
<td>1.506788</td>
</tr>
<tr>
<td>Maximum</td>
<td>147.41</td>
<td>1.85325</td>
<td>1.6147</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>10.00708</td>
<td>0.119753</td>
<td>0.188313</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.254427</td>
<td>-0.723232</td>
<td>0.295178</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.435041</td>
<td>-0.2974507</td>
<td>1.273715</td>
</tr>
</tbody>
</table>

Panel 2: Daily Return Data

<table>
<thead>
<tr>
<th></th>
<th>JPY</th>
<th>SGD</th>
<th>CAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>-0.04565</td>
<td>-0.03523</td>
<td>-0.03737</td>
</tr>
<tr>
<td>1st Quantile</td>
<td>-0.00384</td>
<td>-0.00184</td>
<td>-0.00301</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.00012</td>
<td>-5.9E-05</td>
<td>-9.3E-05</td>
</tr>
<tr>
<td>Median</td>
<td>0</td>
<td>-7.2E-05</td>
<td>-5E-05</td>
</tr>
<tr>
<td>3rd Quantile</td>
<td>0.004003</td>
<td>0.001681</td>
<td>0.002774</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.049073</td>
<td>0.031479</td>
<td>0.033091</td>
</tr>
<tr>
<td></td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.007231</td>
<td>0.003634</td>
<td>0.005528</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.43005</td>
<td>-0.20578</td>
<td>0.043318</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.67931</td>
<td>9.871163</td>
<td>4.425934</td>
</tr>
</tbody>
</table>
Table 5: Backtesting Outcomes of Exchange Rate VaR Estimates

<table>
<thead>
<tr>
<th>Country</th>
<th>RiskMetrics</th>
<th>JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Significance</td>
<td>1%</td>
</tr>
<tr>
<td>JPY</td>
<td>LRuc</td>
<td>Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRand</td>
<td>Don’t Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRcc</td>
<td>Reject VaR Model</td>
</tr>
<tr>
<td>SGD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>1%</td>
</tr>
<tr>
<td>JPY</td>
<td>LRuc</td>
<td>Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRand</td>
<td>Don’t Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRcc</td>
<td>Reject VaR Model</td>
</tr>
<tr>
<td>CAD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>1%</td>
</tr>
<tr>
<td>JPY</td>
<td>LRuc</td>
<td>Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRand</td>
<td>Don’t Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRcc</td>
<td>Reject VaR Model</td>
</tr>
<tr>
<td>SV</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>1%</td>
</tr>
<tr>
<td>JPY</td>
<td>LRuc</td>
<td>Don’t Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>LRand</td>
<td>Don’t Reject VaR Model</td>
</tr>
<tr>
<td></td>
<td>SGD</td>
<td>CAD</td>
</tr>
<tr>
<td>------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td><strong>Significance</strong></td>
<td><strong>1%</strong></td>
<td><strong>5%</strong></td>
</tr>
<tr>
<td>LRuc</td>
<td>Don't Reject VaR Model</td>
<td>LRuc</td>
</tr>
<tr>
<td>LRind</td>
<td>Don't Reject VaR Model</td>
<td>LRind</td>
</tr>
<tr>
<td>LRcc</td>
<td>Don't Reject VaR Model</td>
<td>LRcc</td>
</tr>
</tbody>
</table>
Table 6: Descriptive Data for S&P500 & VIX

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>676.53</td>
<td>9.89</td>
<td>-0.0947</td>
<td>-0.29987</td>
</tr>
<tr>
<td>1st Quantile</td>
<td>1190.47</td>
<td>12.26</td>
<td>-0.00532</td>
<td>-0.03683</td>
</tr>
<tr>
<td>Mean</td>
<td>1248.455</td>
<td>21.26077</td>
<td>-0.00018</td>
<td>0.000435</td>
</tr>
<tr>
<td>Median</td>
<td>1275.55</td>
<td>15.59</td>
<td>0.000739</td>
<td>-0.0033</td>
</tr>
<tr>
<td>3rd Quantile</td>
<td>1400.565</td>
<td>25.140</td>
<td>0.005524</td>
<td>0.032046</td>
</tr>
<tr>
<td>Maximum</td>
<td>1565.15</td>
<td>80.86</td>
<td>0.102457</td>
<td>0.496008</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>203.4692</td>
<td>13.0195</td>
<td>0.015136</td>
<td>0.066551</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.85291</td>
<td>1.859021</td>
<td>-0.10722</td>
<td>0.574778</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.02393</td>
<td>3.345993</td>
<td>9.3311414.770059</td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td>1%</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRuc</td>
<td>Reject VaR Model</td>
<td>Reject VaR Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRind</td>
<td>Reject VaR Model</td>
<td>Don't Reject VaR Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRec</td>
<td>Reject VaR Model</td>
<td>Reject VaR Model</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Historical Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significance</td>
</tr>
<tr>
<td>LRuc</td>
</tr>
<tr>
<td>LRind</td>
</tr>
<tr>
<td>LRec</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significance</td>
</tr>
<tr>
<td>LRuc</td>
</tr>
<tr>
<td>LRind</td>
</tr>
<tr>
<td>LRec</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significance</td>
</tr>
<tr>
<td>LRuc</td>
</tr>
<tr>
<td>LRind</td>
</tr>
<tr>
<td>LRec</td>
</tr>
</tbody>
</table>
Figure 1: Time Series Plots of Exchange Rate Series

Panel 1: Original Daily Data
Panel 2: Daily Return Series

JPY

SGD

CAD
Figure 2: Parameter Estimates of the Improved Estimation Procedure for Exchange Rate Series

Panel 1: JPY

Panel 2: SGD

Panel 3: CAD
Figure 3: Time Series Plot of VaR and CVaR Estimates by the Proposed Procedure for Foreign Exchange Rate

Panel 1: JPY
Panel 2:
SGD

![Graph showing the performance of SGD, VaR99, and CVaR99 over years 2000 to 2009.](image-url)

- SGD
- VaR99
- CVaR99

The graph illustrates the fluctuation of these values over the years, with SGD showing a consistent performance pattern.
Panel 3:
CAD

![Graph of CAD VaR95 and CVaR95 from 2000 to 2009]
Figure 4: Time Series Plots of S&P 500 and Its VIX

Panel 1: Original Daily Data

Panel 2: Daily Return Series
Figure 5: Time Series Plot of Rho between S&P500 and VIX

Figure 6: Parameter Estimates of the Improved Estimation Procedure for S&P 500 with VIX Series
Figure 7: Time Series Plot of VaR and CVaR Estimates by the Proposed Procedure for S&P 500
References


